Online Supplement

Supplement A: Auxiliary Results and Proofs

Lemma 1. *The function $G_{\mathcal{S}}$ defined in (5) is a continuous, symmetric and positive definite function. Moreover it is a valid covariance function.*

Proof. Symmetry is obvious. Continuity follows from the continuity of the covariance function C. We have that $G_{\mathcal{S}}(s,u) = \int_{\mathcal{T}} \text{cov}(X(s,t), X(u,t))dt$. Observing that for each fixed t, $\text{cov}(X(s,t), X(u,t))$ is positive definite, it holds that for any function f in $L^2(\mathcal{S})$ we have that $\int_{\mathcal{S}\times\mathcal{S}} \text{cov}(X(s,t), X(u,t))f(u)f(s)duds \geq 0$ for all $t \in \mathcal{T}$. Therefore, by Fubini,

$$\int_{\mathcal{S}\times\mathcal{S}} G_{\mathcal{S}}(u,s)f(u)f(s)duds = \int_{\mathcal{T}}\int_{\mathcal{S}\times\mathcal{S}} \text{cov}(X(s,t), X(u,t))f(u)f(s)dudsdt \geq 0.$$

To see that $G_{\mathcal{S}}$ is a valid covariance function, remember that this is equivalent to say that $\int_{\mathcal{S}} G_{\mathcal{S}}(s,s)ds < \infty$ (see, for instance, Horváth and Kokoszka 2012, page 24). Observe that

$$\int_{\mathcal{S}} G_{\mathcal{S}}(s,s)ds = \int_{\mathcal{S}}\int_{\mathcal{T}} C((s,t),(s,t))dtds$$

and the last integral is finite because C is a valid covariance function. □

Lemma 2. *Let $\xi_j(t) = \langle X(\cdot,t) - \mu(\cdot,t), \psi_j\rangle_{\mathcal{S}}$, $j \geq 1$, be the random functional coefficients in the marginal Karhunen-Loève representation (7). Then $E(\xi_j(t)) = 0$ for almost all $t \in \mathcal{T}$, and $E(\langle \xi_j, \xi_k\rangle_{\mathcal{T}}) = \tau_j\delta_{jk}$, where $\delta_{jk} = 1$ if $j = k$ and $= 0$ otherwise.*

Proof. First, for almost all $t \in \mathcal{T}$, $X(\cdot,t)$ is a random element of $L^2(\mathcal{S})$ because X is in $L^2(\mathcal{S}\times\mathcal{T})$. Then there exists a unique (in the L^2 sense) function $\tilde{\mu}(\cdot,t)$ in $L^2(\mathcal{S})$ such that $E(\langle X(\cdot,t), y\rangle_{\mathcal{S}}) = \langle \tilde{\mu}(\cdot,t), y\rangle_{\mathcal{S}}$ for all $y \in L^2(\mathcal{S})$ and it follows that $\tilde{\mu}(s,t) = E(X(s,t)) = \mu(s,t)$ for almost all $s \in \mathcal{S}$ (see, for instance, Horváth and Kokoszka (2012), Section 2.3), so that $\tilde{\mu}(\cdot,t) = \mu(\cdot,t)$ in the L^2 sense. Then taking $y = \psi_j$ we have that

$$E(\xi_j(t)) = E\left(\langle X(\cdot,t), \psi_j\rangle_{\mathcal{S}}\right) - \langle \mu(\cdot,t), \psi_j\rangle_{\mathcal{S}} = 0.$$

Furthermore,

$$
\begin{aligned}
&E(\langle \xi_j, \xi_k \rangle_{\mathcal{T}}) = E\left(\int_{\mathcal{T}} \xi_j(t)\xi_k(t)dt \right) \\
=&E\left(\int_{\mathcal{T}} \left(\int_{\mathcal{S}} X^c(s,t)\psi_j(s)ds \right) \left(\int_{\mathcal{S}} X^c(u,t)\psi_k(u)du \right) dt \right) \\
=&\int_{\mathcal{S}}\int_{\mathcal{S}}\int_{\mathcal{T}} E\left(X^c(s,t)X^c(u,t)\right) dt\ \psi_j(s)\psi_k(u)\ ds\ du \\
=&\int_{\mathcal{S}} \left(\int_{\mathcal{S}} G_{\mathcal{S}}(s,u)\psi_j(s)\ ds \right) \psi_k(u)\ du \\
=&\langle \Psi(\psi_j), \psi_k \rangle_{\mathcal{S}} = \tau_j \langle \psi_j, \psi_k \rangle_{\mathcal{S}} = \tau_j \delta_{jk},
\end{aligned}
$$

where we have used that τ_j, ψ_j, $j \geq 1$, are, respectively, the eigenvalues and eigenfunctions of the operator Ψ defined as $\Psi(f)(s) = \int G_{\mathcal{S}}(s,u)\, f(u)\, du$. □

Proof of Theorem 1:

Observe that

$$
\begin{aligned}
&E\left(\int_{\mathcal{T}} \|X^c(\cdot,t) - \sum_{j=1}^{P} \langle X^c(\cdot,t), g_j \rangle_{\mathcal{S}} g_j \|_{\mathcal{S}}^2 dt \right) \\
=&E\left(\int_{\mathcal{T}} \langle X^c(\cdot,t) - \sum_{j=1}^{P} \langle X^c(\cdot,t), g_j \rangle_{\mathcal{S}} g_j, X^c(\cdot,t) - \sum_{j=1}^{P} \langle X^c(\cdot,t), g_j \rangle_{\mathcal{S}} g_j \rangle_{\mathcal{S}} dt \right) \\
=&E\left(\int_{\mathcal{T}} \langle X^c(\cdot,t), X^c(\cdot,t) \rangle_{\mathcal{S}} dt \right) - \sum_{j=1}^{P} E\left(\int_{\mathcal{T}} \langle X^c(\cdot,t), g_j \rangle_{\mathcal{S}}^2 dt \right) \\
=&E\left(\int_{\mathcal{T}} \int_{\mathcal{S}} \left(X^c(s,t)\right)^2 ds dt \right) - \sum_{j=1}^{P} E\left(\int_{\mathcal{T}} \left(\int_{\mathcal{S}} X^c(s,t) g_j(s) ds \right)^2 dt \right) \\
=&E\left(\|X^c\|^2 \right) - \sum_{j=1}^{P} \int_{\mathcal{S}}\int_{\mathcal{S}}\int_{\mathcal{T}} E\left(X^c(s,t)X^c(u,t)\right) dt\ g_j(u) du\ g_j(s) ds \\
=&E\left(\|X^c\|^2 \right) - \sum_{j=1}^{P} \int_{\mathcal{S}}\int_{\mathcal{S}}\int_{\mathcal{T}} C((s,t),(u,t)) dt\ g_j(u) du\ g_j(s) ds \\
=&E\left(\|X^c\|^2 \right) - \sum_{j=1}^{P} \int_{\mathcal{S}}\int_{\mathcal{S}} G_{\mathcal{S}}(s,u) g_j(u) du\ g_j(s) ds \\
=&E\left(\|X^c\|^2 \right) - \sum_{j=1}^{P} \langle \Psi(g_j), g_j \rangle_{\mathcal{S}},
\end{aligned}
$$

where Ψ is the operator in $L^2(\mathcal{S})$ with kernel $G_{\mathcal{S}}$. Then minimizing

$$E\left(\int_{\mathcal{T}} \|X^c(\cdot,t) - \sum_{j=1}^{P} \langle X^c(\cdot,t), g_j\rangle_{\mathcal{S}} g_j\|_{\mathcal{S}}^2 dt\right)$$

is equivalent to maximizing $\sum_{j=1}^{P} \int_{\mathcal{S}} \langle \Psi(g_j), g_j\rangle_{\mathcal{S}}$. Given that Ψ is a symmetric, positive definite Hilbert-Schmidt operator (see Lemma 1), standard arguments (see, for instance, Theorem 3.2 in Horváth and Kokoszka (2012)) complete the proof.

Proof of Theorem 2:

For $f_{jk}(t)$ and $g_j(s)$ that satisfy the orthogonal conditions, we have

$$\begin{aligned}
&E\left(\int_{\mathcal{S},\mathcal{T}} \{X^c(s,t) - \sum_{j=1}^{P}\sum_{k=1}^{K_j} \langle X^c, f_{jk}g_j\rangle f_{jk}(t)g_j(s)\}^2 dsdt\right) \\
=&E\|X^c\|^2 - 2\times \sum_{j=1}^{P}\sum_{k=1}^{K_j} E\left(\int_{\mathcal{S},\mathcal{T}} X^c(s,t)\langle X^c, f_{jk}g_j\rangle f_{jk}(t)g_j(s)\right) dsdt \\
&+ E\left(\int_{\mathcal{S},\mathcal{T}} \sum_{j=1}^{P}\sum_{k=1}^{K_j}\sum_{l=1}^{P}\sum_{h=1}^{K_j} \langle X^c, f_{jk}g_j\rangle f_{jk}(t)g_j(s)\langle X^c, f_{lh}g_l\rangle f_{lh}(t)g_l(s) dsdt\right) \\
=&E\|X^c\|^2 - 2\times \sum_{j=1}^{P}\sum_{k=1}^{K_j} \int E\left(X^c(s,t)X^c(u,v)\right) f_{jk}(t)f_{jk}(v)g_j(s)g_j(u) dsdtdudv \\
&+ \sum_{j=1}^{P}\sum_{k=1}^{K_j} \int E\left(X^c(s,t)X^c(u,v)\right) f_{jk}(t)f_{jk}(v)g_j(s)g_j(u) dsdtdudv \\
=&E\|X^c\|^2 - \sum_{j=1}^{P}\sum_{k=1}^{K_j} \int E\left(X^c(s,t)X^c(u,v)\right) f_{jk}(t)f_{jk}(v)g_j(s)g_j(u) dsdtdudv
\end{aligned} \tag{24}$$

Let $\tilde{f}_{jk}(t)$ and $\tilde{g}_j(s)$ denote the optimal basis to achieve the minimum reconstruction error Q^*, and define

$$(I) = \sum_{j=1}^{P}\sum_{k=1}^{K_j} \int E\left(X^c(s,t)X^c(u,v)\right) \phi_{jk}(t)\phi_{jk}(v)\psi_j(s)\psi_j(u) dsdtdudv$$

and

$$(II) = \sum_{j=1}^{P}\sum_{k=1}^{K_j} \int E\left(X^c(s,t)X^c(u,v)\right) \tilde{f}_{jk}(t)\tilde{f}_{jk}(v)\tilde{g}_j(s)\tilde{g}_j(u) dsdtdudv.$$

By Eq. (24), to prove the theorem, we only need to show that $(II) - (I) < aE\|X^c\|_2$.

We further define

$$(III) = \sum_{j=1}^{P} \int_{\mathcal{S}\times\mathcal{S}} \int_{\mathcal{T}} E\left(X^c(s,t)X^c(u,t)\right) dt\tilde{g}_j(s)\tilde{g}_j(u)dsdu,$$

and

$$(IV) = \sum_{j=1}^{P} \int_{\mathcal{S}\times\mathcal{S}} \int_{\mathcal{T}} E\left(X^c(s,t)X^c(u,t)\right) dt\psi_j(s)\psi_j(u)dsdu.$$

We will prove that $(II) < (III) < (IV)$ and $(IV) - (I) < aE\|X^c\|_2$, which implies that $(II) - (I) < aE\|X^c\|_2$.

By definition, ψ_j are the leading eigenfunctions of the marginal kernel $G_{\mathcal{S}}(s,u)$ so that $(III) < (IV)$.

To show $(II) < (III)$, let $\tilde{\xi}_j(t) = \langle X^c(s,t), \tilde{g}_j(s)\rangle$. Then,

$$\begin{aligned}
(II) =& \sum_{j=1}^{P}\sum_{k=1}^{K_j} \int E\left(X^c(s,t)X^c(u,v)\right) \tilde{f}_{jk}(t)\tilde{f}_{jk}(v)\tilde{g}_j(s)\tilde{g}_j(u)dsdtdudv \\
=& \sum_{j=1}^{P}\sum_{k=1}^{K_j} \int E\left(\tilde{\xi}_j(t)\tilde{\xi}_j(v)\right) \tilde{f}_{jk}(t)\tilde{f}_{jk}(v)dtdv \\
<& \sum_{j=1}^{P}\sum_{k=1}^{\infty} \int E\left(\tilde{\xi}_j(t)\tilde{\xi}_j(v)\right) \tilde{f}_{jk}(t)\tilde{f}_{jk}(v)dtdv \\
=& \sum_{j=1}^{P} \int E\left(\tilde{\xi}_j(t)\tilde{\xi}_j(t)\right) dt \\
=& \sum_{j=1}^{P} \int E\left(\int X^c(s,t)\tilde{g}_j(s)ds \int X^c(u,t)\tilde{g}_j(u)du\right) dt \\
=& (III)
\end{aligned} \tag{25}$$

Finally, we prove $(IV) - (I) < aE\|X^c\|_2$,

$$\begin{aligned}(IV) - (I) =& \sum_{j=1}^{P} \int E\left(\xi_j(t)\xi_j(t)\right) dt \\ & - \sum_{j=1}^{P}\sum_{k=1}^{K_j} \int E\left(X^c(s,t)X^c(u,v)\right)\phi_{jk}(t)\phi_{jk}(v)\psi_j(s)\psi_j(u)dsdtdudv \\ =& \sum_{j=1}^{P}\sum_{k=1}^{\infty} \int E\left(\xi_j(t)\xi_j(v)\right)\phi_{jk}(t)\phi_{jk}(v)dtdv \\ & - \sum_{j=1}^{P}\sum_{k=1}^{K_j} \int E\left(\xi_j(t)\xi_j(v)\right)\phi_{jk}(t)\phi_{jk}(v)dtdv \\ & < aE\|X^c\|_2,\end{aligned} \tag{26}$$

where $a = \max_{1\le j\le P} a_j$, with $(1-a_j)\%$ denoting the fraction of variance explained by K_j terms in each process $\xi_j(t) = \langle X^c(\cdot,t), \psi_j\rangle$.

Proof of Theorem 3:

Recall that

$$G_{\mathcal{S}}(s,u) = \int_{\mathcal{T}} C((s,t),(u,t))dt,$$

where $C((s,t),(u,t)) = E[(X(s,t)-\mu(s,t))(X(u,t)-\mu(u,t))]$. For (s,u) on the grid points, we have

$$\hat{G}_{\mathcal{S}}(s,u) = \frac{|\mathcal{T}|}{\sum_{i=1}^{n} M_i}\sum_{i=1}^{n}\sum_{m=1}^{M_i} \hat{X}^c(s,t_{im})\hat{X}^c(u,t_{im}),$$

where $\hat{X}^c(s,t_{im}) = X(s,t_{im}) - \hat{\mu}(s,t_{im})$ and $|\mathcal{T}|$ is the Lebesgue measure of $\mathcal{T}$. We define

$$\tilde{G}_{\mathcal{S}}(s,u) = \frac{|\mathcal{T}|}{\sum_{i=1}^{n} M_i}\sum_{i=1}^{n}\sum_{m=1}^{M_i} X^c(s,t_{im})X^c(u,t_{im}).$$

Using $\sup_{s,t}|\hat{\mu}(s,t) - \mu(s,t)| = O_p((\log n/n)^{1/2})$, it is easy to show that $\|\tilde{G}_{\mathcal{S}}(s,u) - \hat{G}_{\mathcal{S}}(s,u)\|_{\mathcal{S}} = O_p((\log n/n)^{1/2})$. Next we show

$$\|\tilde{G}_{\mathcal{S}}(s,u) - G_{\mathcal{S}}(s,u)\|_{\mathcal{S}} = O_p((1/n)^{1/2}). \tag{27}$$

We first prove (27) under assumption (A.6a). By (A.6a), we have $M_i \equiv M$, and the grid of t is $\{t_1, \ldots, t_M\}$. Therefore,

$$\begin{aligned} &\sup_{(s,u)\in\mathcal{S}^2} |E\tilde{G}_{\mathcal{S}}(s,u) - G_{\mathcal{S}}(s,u)| \\ &= \sup_{(s,u)\in\mathcal{S}^2} \left| \frac{|\mathcal{T}|}{M} \sum_{m=1}^{M} C((s,t_m),(u,t_m)) - \int_{\mathcal{T}} C((s,t),(u,t))dt \right| \\ &\overset{A.3}{=} O(\frac{1}{M}) = O(1/n), \end{aligned} \tag{28}$$

and

$$\begin{aligned} &\sup_{(s,u)\in\mathcal{S}^2} \mathrm{var}(\tilde{G}_{\mathcal{S}}(s,u)) \\ &= \sup_{(s,u)\in\mathcal{S}^2} \frac{|\mathcal{T}|^2}{(nM)^2} \sum_{i=1}^{n} \mathrm{var}(\sum_{m=1}^{M} X^c(s,t_m)X^c(u,t_m)) \\ &\leq \sup_{(s,u)\in\mathcal{S}^2} \frac{|\mathcal{T}|^2}{(nM)^2} \sum_{i=1}^{n} \sum_{m,m'=1}^{M} E(X^c(s,t_m)X^c(u,t_m)X^c(s,t_{m'}))X^c(u,t_{m'})) \\ &\overset{A.1,A.2}{\leq} \frac{|\mathcal{T}|^2}{n^2M^2} \sum_{i=1}^{n} M^2B = O(1/n). \end{aligned} \tag{29}$$

Combining (28) and (29) we have

$$\sup_{(s,u)\in\mathcal{S}^2} E|\tilde{G}_{\mathcal{S}}(s,u) - G_{\mathcal{S}}(s,u)|^2 = O(1/n).$$

Therefore, by (A.4) and $(s_l - s_{l-1}) = O(n^{-1})$,

$$\begin{aligned} E\|\tilde{G}_{\mathcal{S}}(s,u) - G_{\mathcal{S}}(s,u)\|_{\mathcal{S}}^2 &= \int_{\mathcal{S}}\int_{\mathcal{S}} |\tilde{G}_{\mathcal{S}}(s,u) - G_{\mathcal{S}}(s,u)|^2 dsdu \\ &= \frac{|\mathcal{S}|}{L^2} \sum_{j=1}^{L}\sum_{l=1}^{L} E|\tilde{G}_{\mathcal{S}}(s_j,s_l) - G_{\mathcal{S}}(s_j,s_l)|^2 + O(1/n) \\ &= O(1/n), \end{aligned} \tag{30}$$

which implies (27). The same result can be derived using a similar argument under (A.6b),

by substituting (28) with

$$\sup_{(s,u)\in\mathcal{S}^2} |E\tilde{G}_{\mathcal{S}}(s,u) - G_{\mathcal{S}}(s,u)| \tag{31}$$
$$= \left| E\frac{|\mathcal{T}|}{\sum_{i=1}^n M_i} \sum_{i=1}^n \sum_{m=1}^{M_i} C((s,t_{im}),(u,t_{im})) - \int_{\mathcal{T}} C((s,t),(u,t))dt \right| = 0,$$

and substituting (29) with

$$\sup_{(s,u)\in\mathcal{S}^2} \mathrm{var}(\tilde{G}_{\mathcal{S}}(s,u)) = \frac{|\mathcal{T}|^2}{(\sum_{i=1}^n M_i)^2} \sum_{i=1}^n \mathrm{var}(\sum_{m=1}^{M_i} X^c(s,t_{im})X^c(u,t_{im})) \tag{32}$$
$$\overset{A.1,A.2}{\leq} \frac{|\mathcal{T}|^2}{(\sum_{i=1}^n M_i)^2} \sum_{i=1}^n M_i^2 B = O(1/n).$$

This completes the proof for Eq. (15).

For a fixed j, Lemma 4.3 in Bosq (2000) implies that

$$|\hat{\tau}_j - \tau_j| \leq ||\hat{G}_{\mathcal{S}}(s,u) - G_{\mathcal{S}}(s,u)||_{\mathcal{S}}, \quad ||\hat{\psi}_j(s) - \psi_j(s)||_{\mathcal{S}} \leq 2\sqrt{2}\delta_j^{-1}||\hat{G}_{\mathcal{S}}(s,u) - G_{\mathcal{S}}(s,u)||, \tag{33}$$

where δ_j is defined in (A.5). By (A.5), Eq. (16) and Eq. (17) directly follow.

In the following, we establish Eq. (18) as follows,

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^n \sup_{1\leq m\leq M_i} |\hat{\xi}_{ij}(t_{im}) - \xi_{ij}(t_{im})| \\
\leq &\frac{1}{n}\sum_{i=1}^n \sup_{1\leq m\leq M_i} |\int (X_i(s,t_{im}) - \mu(s,t_{im}))(\psi_j(s) - \hat{\psi}_j(s))ds| \\
&+ \frac{1}{n}\sum_{i=1}^n \sup_{1\leq m\leq M_i} |\int (\mu(s,t_{im}) - \hat{\mu}(s,t_{im}))\psi_j(s)ds| \\
&+ \frac{1}{n}\sum_{i=1}^n \sup_{1\leq m\leq M_i} |\int (\hat{\mu}(s,t_{im}) - \mu(s,t_{im}))(\psi_j(s) - \hat{\psi}_j(s))ds| \\
\leq &\frac{1}{n}\sum_{i=1}^n \sup_{s,t} |X_i(s,t_{im}) - \mu(s,t)| \|\psi_j(s) - \hat{\psi}_j(s)\|_{\mathcal{S}} \\
&+ \sup_{s,t} |\mu(s,t) - \hat{\mu}(s,t)| \sup_s |\psi_j(s)| + \sup_{s,t} |\hat{\mu}(s,t) - \mu(s,t)| \|\psi_j(s) - \hat{\psi}_j(s)\|_{\mathcal{S}} \\
&= O_p((\log n/n)^{1/2}) + O_p((\log n/n)^{1/2}) + O_p((\log n/n)^{1/2}) \\
&= O_p((\log n/n)^{1/2}),
\end{aligned} \tag{34}$$

where we used (A.1),(A.2), $\sup_{s,t}|\hat{\mu}(s,t)-\mu(s,t)| = O_p((\log n/n)^{1/2})$, and the previous result $\sup_s|\hat{\phi}_j(s)-\phi_j(s)| = O_p((\log n/n)^{1/2})$. This completes the proof.

Proof of Theorem 4:

$$\begin{aligned}
&C((s,t),(u,v)) = \mathrm{cov}(X(s,t),X(u,v))\\
=&\mathrm{cov}\left(\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\chi_{jk}f_k(t)g_j(s),\sum_{l=1}^{\infty}\sum_{h=1}^{\infty}\chi_{lh}f_h(v)g_l(u)\right)\\
=&\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\sum_{h=1}^{\infty}\mathrm{cov}(\chi_{jk}\chi_{lh})f_k(t)g_j(s)f_h(v)g_l(u).
\end{aligned}$$

Furthermore, by the orthogonality of f_k and $\mathrm{cov}(\chi_{jk},\chi_{lk}) = 0$ for $j \neq l$, we have

$$C_{\mathcal{S}}(s,u) = \int_{\mathcal{T}} C((s,t),(u,t))dt = \sum_{j=1}^{\infty}\sum_{l=1}^{\infty}\sum_{k=1}^{\infty}\sum_{h=1}^{\infty}\mathrm{cov}(\chi_{jk}\chi_{lh})g_j(s)g_l(u)\int_{\mathcal{T}} f_k(t)f_h(t)dt$$

$$=\sum_{j=1}^{\infty}(\sum_{k=1}^{\infty}\mathrm{cov}(\chi_{jk},\chi_{jk}))g_j(s)g_j(u).$$

Therefore, $g_j(s)$ are the unique eigenfunctions of $C_{\mathcal{S}}(s,u)$, and $\tau_j = \sum_{k=1}^{\infty}\mathrm{var}(\chi_{jk})$. By symmetry, one obtains the analogous result $f_k(t) \equiv \phi_k(t)$.

Proof of Theorem 5:

For $f_k(t)$ and $g_j(s)$ that satisfy the orthogonality conditions,

$$\begin{aligned}
&E\left(\int_{\mathcal{S},\mathcal{T}}\{X^c(s,t)-\sum_{j=1}^{P}\sum_{k=1}^{K}\langle X^c,f_kg_j\rangle f_k(t)g_j(s)\}^2dsdt\right)\\
=&E\|X^c\|^2 - 2\times\sum_{j=1}^{P}\sum_{k=1}^{K}\int E\left(X^c(s,t)X^c(u,v)\right)f_k(t)f_k(v)g_j(s)g_j(u)dsdtdudv\\
&+\sum_{j=1}^{P}\sum_{k=1}^{K}\int E\left(X^c(s,t)X^c(u,v)\right)f_k(t)f_k(v)g_j(s)g_j(u)dsdtdudv\\
=&E\|X^c\|^2 - \sum_{j=1}^{P}\sum_{k=1}^{K}\int E\left(X^c(s,t)X^c(u,v)\right)f_k(t)f_k(v)g_j(s)g_j(u)dsdtdudv.
\end{aligned} \tag{35}$$

Let $\tilde{f}_k(t)$ and $\tilde{g}_j(s)$ denote the optimal basis to achieve the minimum reconstruction error Q^*, and define

$$(I) = \sum_{j=1}^{P}\sum_{k=1}^{K}\int E\left(X^c(s,t)X^c(u,v)\right)\phi_k(t)\phi_k(v)\psi_j(s)\psi_j(u)dsdtdudv,$$

and

$$(II) = \sum_{j=1}^{P}\sum_{k=1}^{K}\int E\left(X^c(s,t)X^c(u,v)\right)\tilde{f}_k(t)\tilde{f}_k(v)\tilde{g}_j(s)\tilde{g}_j(u)dsdtdudv.$$

By Eq. (35), to prove the theorem, we only need to show that $(II) - (I) < aE\|X^c\|_2$.

We further define

$$(III) = \sum_{j=1}^{P}\int_{\mathcal{S}\times\mathcal{S}}\int_{\mathcal{T}} E\left(X^c(s,t)X^c(u,t)\right)dt\tilde{g}_j(s)\tilde{g}_j(u)dsdu,$$

and

$$(IV) = \sum_{j=1}^{P}\int_{\mathcal{S}\times\mathcal{S}}\int_{\mathcal{T}} E\left(X^c(s,t)X^c(u,t)\right)dt\psi_j(s)\psi_j(u)dsdu.$$

We will prove that $(II) < (III) < (IV)$ and $(IV) - (I) < aE\|X^c\|_2$, which implies that $(II) - (I) < aE\|X^c\|_2$.

By definition, the ψ_j are the leading eigenfunctions of the marginal kernel $G_{\mathcal{S}}(s,u)$ so that $(III) < (IV)$.

To show $(II) < (III)$, let $\tilde{\xi}_j(t) = \langle X^c(s,t), \tilde{g}_j(s)\rangle$, we observe

$$\begin{aligned}
(II) =& \sum_{j=1}^{P}\sum_{k=1}^{K}\int E\left(X^c(s,t)X^c(u,v)\right)\tilde{f}_k(t)\tilde{f}_k(v)\tilde{g}_j(s)\tilde{g}_j(u)dsdtdudv \\
=& \sum_{j=1}^{P}\sum_{k=1}^{K}\int E\left(\tilde{\xi}_j(t)\tilde{\xi}_j(v)\right)\tilde{f}_k(t)\tilde{f}_k(v)dtdv \\
<& \sum_{j=1}^{P}\int E\left(\tilde{\xi}_j(t)\tilde{\xi}_j(t)\right)dt \\
=& \sum_{j=1}^{P}\int E\left(\int X^c(s,t)\tilde{g}_j(s)ds\int X^c(u,t)\tilde{g}_j(u)du\right)dt \\
=& (III). \qquad (36)
\end{aligned}$$

Finally, we prove $(IV) - (I) < aE\|X^c\|_2 = \max(a_{\mathcal{S}}, a_{\mathcal{T}})E\|X^c\|_2$. Recall that τ_j and ϑ_k are the eigenvalues of $G_{\mathcal{T}}(s,u)$ and $G_{\mathcal{S}}(t,v)$. Then

$$\begin{aligned}
\tau_j &= \int G_{\mathcal{T}}(s,u)\psi_j(s)\psi_j(u)dsdu \\
&= \int\int E\left(X^c(s,t)X^c(u,t)\right)dt\psi_j(s)\psi_j(u)dsdu \\
&= \int\int \sum_{j'}\sum_{k}\sum_{l}\sum_{h} E(\chi_{j'k}\chi_{lh})\phi_k(t)\phi_h(t)dt\psi_{j'}(s)\psi_l(u)\psi_j(s)\psi_j(u)dsdu \\
&= \sum_{k=1}^{\infty}\mathrm{var}(\chi_{jk}),
\end{aligned}$$

and by symmetry, we obtain $\vartheta_k = \sum_{j=1}^{\infty}\mathrm{var}(\chi_{jk})$. Then,

$$\begin{aligned}
(IV) - (I) &= \sum_{j=1}^{P}\tau_j - \sum_{j=1}^{P}\sum_{k=1}^{K}\mathrm{var}(\chi_{jk}) = \sum_{j=1}^{P}\sum_{k=K+1}^{\infty}\mathrm{var}(\chi_{jk}) \\
&< \sum_{k=K+1}^{\infty}\vartheta_k = a_{\mathcal{T}}E\|X^c\|_2,
\end{aligned}$$

By symmetry, we also have $(IV) - (I) < a_{\mathcal{S}}E\|X^c\|_2$.

Supplement B: Additional Tables and Figures for the analysis of the fertility data and simulations

Additional materials on the fertility data, which were downloaded from the human fertility database on March 18, 2013, are provided in Table 4 and Figures 5-11. These complement the results presented in sections 5 and 6 of the main part of the paper.

Table 4: The abbreviations and names of the 17 countries (or territories), whose data are used in the fertility application (those with available data for the period 1951-2006). The colors used for representing each country in Figures 4 and 6 are also shown.

Color	Abbreviation	Country name	First year	Last year
	SWE	Sweden	1891	2010
	CAN	Canada	1921	2007
	ESP	Spain	1922	2006
	CHE	Switzerland	1932	2009
	USA	U.S.	1933	2010
	GBRTENW	U.K., England and Wales	1938	2009
	FIN	Finland	1939	2009
	PRT	Portugal	1940	2009
	GBR_SCO	U.K., Scotland	1945	2009
	FRA	France	1946	2010
	BGR	Bulgaria	1947	2009
	JPN	Japan	1947	2009
	CZE	Czech Republic	1950	2011
	HUN	Hungary	1950	2009
	NLD	Netherlands	1950	2009
	SVK	Slovakia	1950	2009
	AUT	Austria	1951	2010

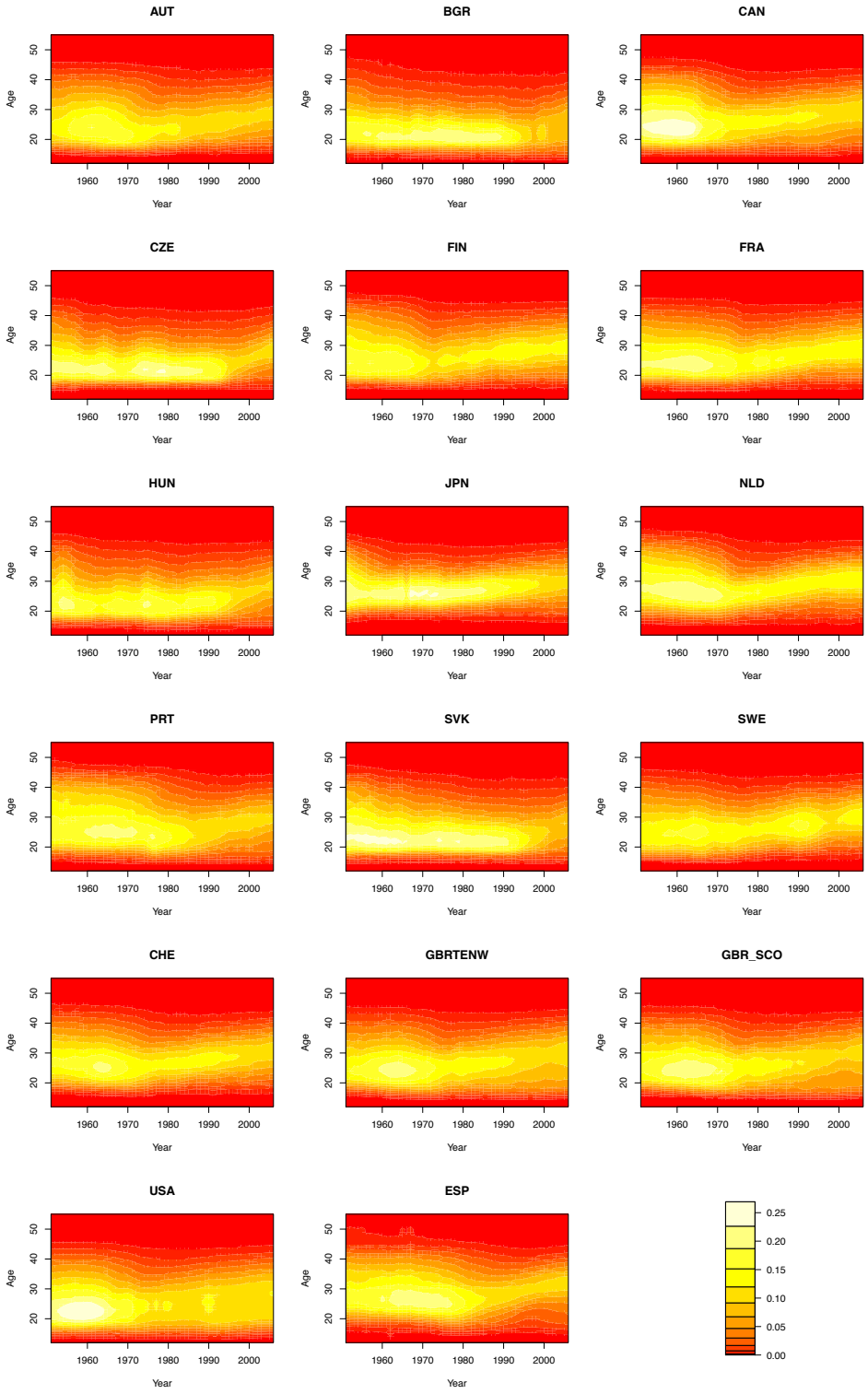

Figure 5: Age-specific fertility rates (ASFR) for 17 countries, red colors correspond to low values and yellow colors to high values.

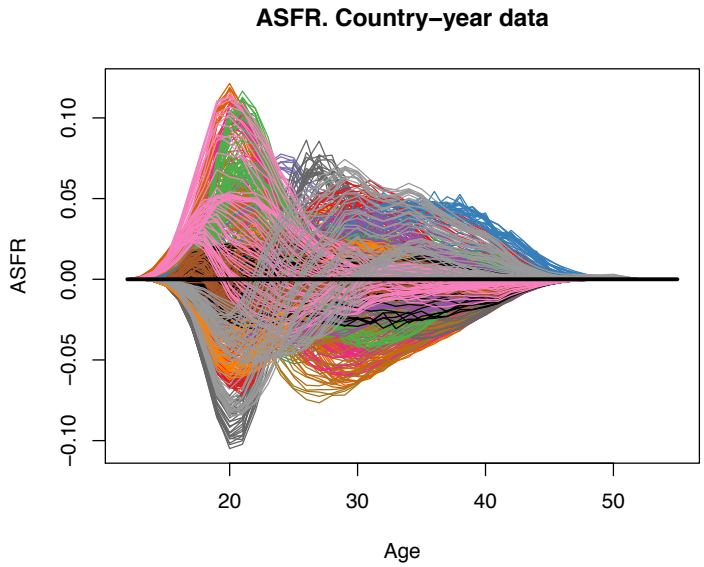

Figure 6: All available functional fertility data as functions of *age* for 952 combinations of 17 countries and 56 calendar years, centered around the mean. Functions corresponding to the same country are in the same color.

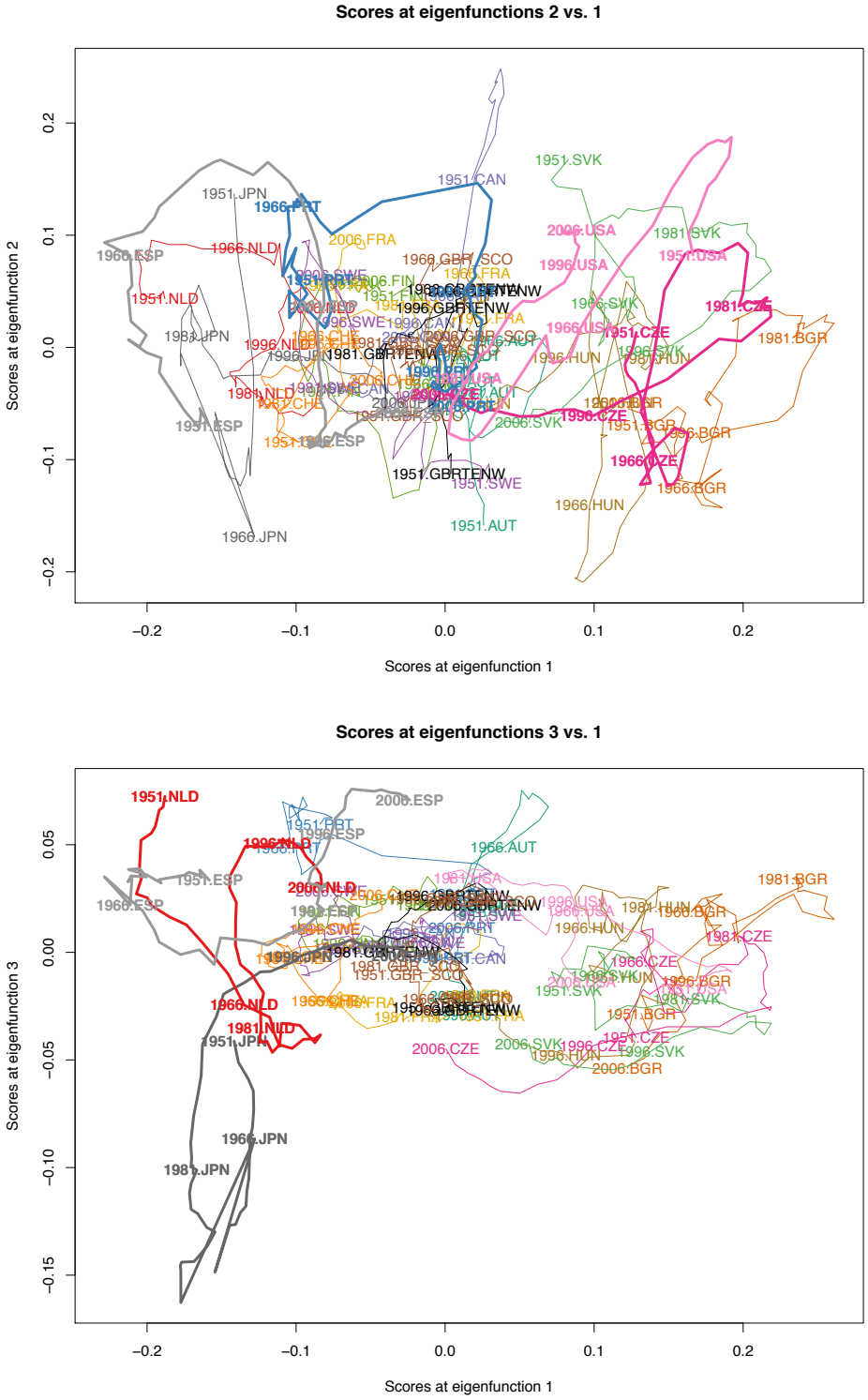

Figure 7: Track-plots corresponding to the implicitly parametrized planar curves $\{(\hat{\xi}_{i,1}(t), \hat{\xi}_{i,2}(t)),\ t = 1951, \ldots, 2006\}$, parametrized by calendar time t, where $\xi_{i,j}(t)$ is the j-th score function for country i as in (4).

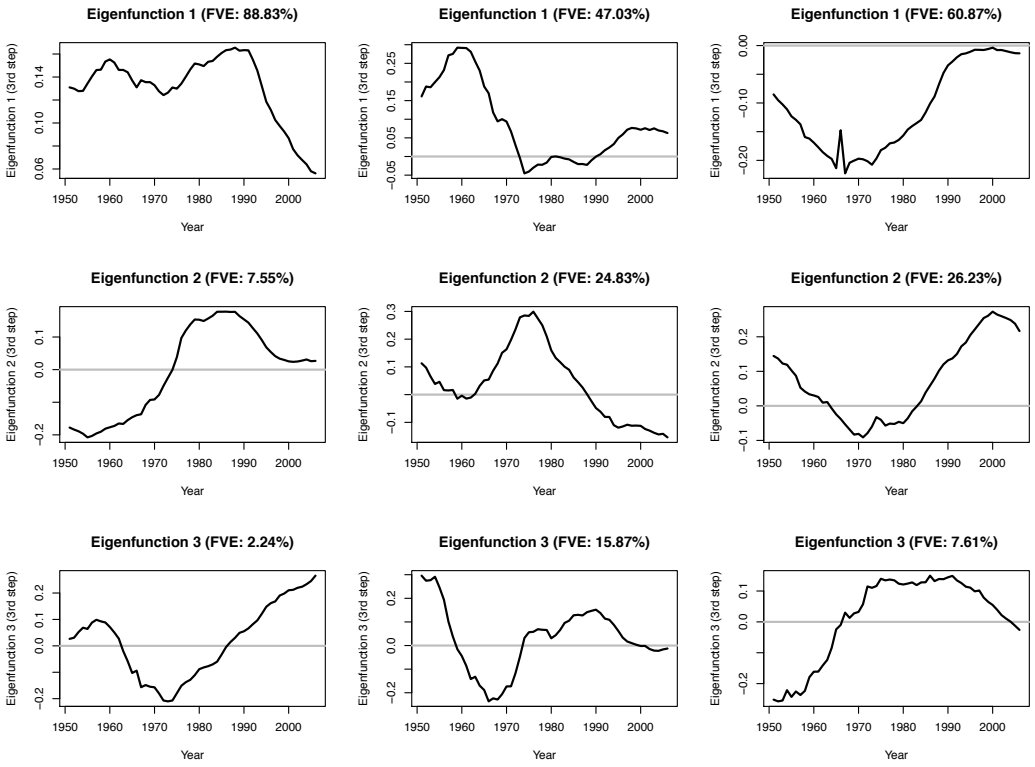

Figure 8: Estimated eigenfunctions $\hat{\phi}_{jk}(t)$ of the random scores $\xi_j(t)$. These quantities are as defined in (4).

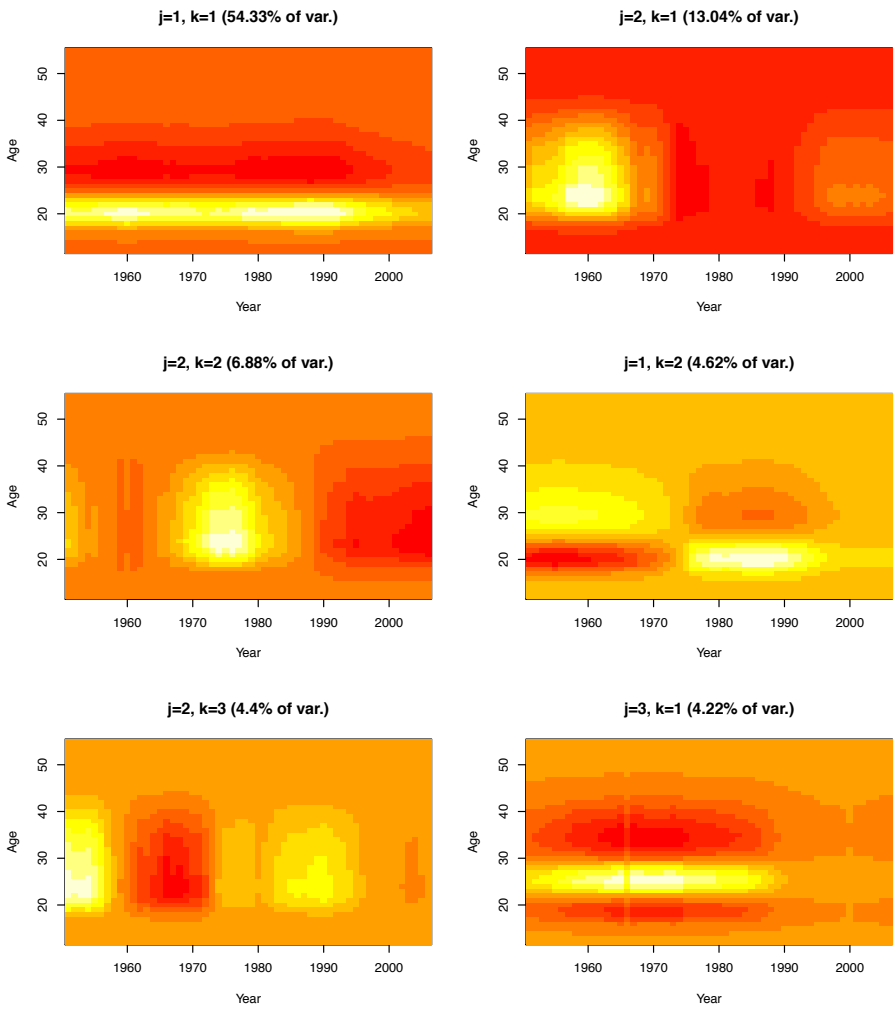

Figure 9: Product functions $\hat{\phi}_{jk}(t)\hat{\psi}_j(s)$ corresponding to the six terms with higher FVE in the marginal FPCA representation (4) of ASFR(s,t).

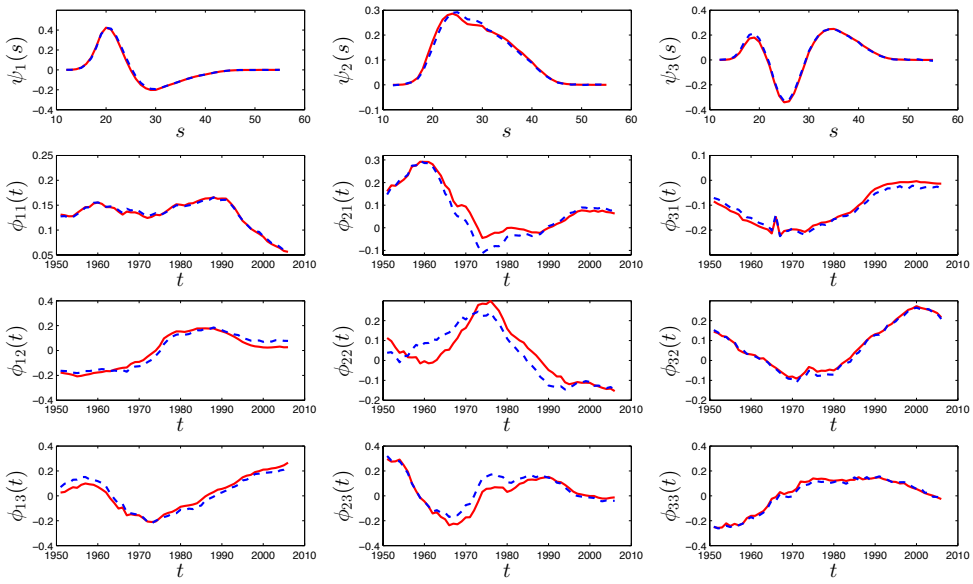

Figure 10: True (red-solid) and estimated (blue-dashed) eigenfunctions $\psi_j(s)$ and $\phi_{jk}(t)$ as in model (4) for $j = 1, 2, 3$ and $k = 1, 2, 3$, for one run of simulation 1 with sample size $n = 50$.

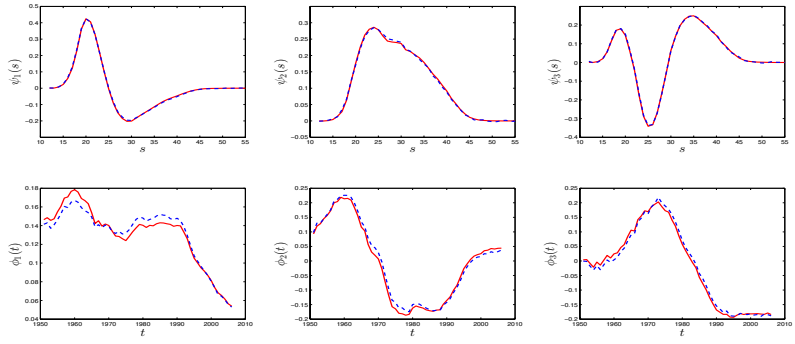

Figure 11: True (red-solid) and estimated (blue-dashed) eigenfunctions $\psi_j(s)$ and $\phi_k(t)$ in model (6) for $j = 1, 2, 3$ and $k = 1, 2, 3$, as obtained in one run of simulation 2 with sample size $n = 50$.

Supplement C: Standard two-dimensional FPCA and product FPCA for fertility data

Here we present the standard FPCA analysis of the ASFR data with the Karhunen-Loève representation, considering the data as random functions in two arguments. We performed FPCA for this type of functional data following Yao *et al.* (2005) as implemented in the PACE package (http://www.stat.ucdavis.edu/PACE). First, we rearrange the $n = 17$ matrices with dimension $L \times M = 44 \times 56$, containing the observed functional data, into a big data matrix with dimension $n \times (M \cdot L)$. Then we perform FPCA on this big matrix. Finally we rearrange the estimated eigenfunctions (stored at this point as arrays of length $M \cdot L$) into matrices of dimension $M \times L$. Figure 12 graphically summarizes the main results of this standard FPCA.

The first 4 eigenfunctions (which are eigensurfaces in this unconstrained approach) have a FVE of 89.73%. The first one (with FVE equal to 58.93%) is almost constant in calendar year and corresponds to a contrast between young fertility (women aged between 18 and 25 years) and fertility in mature years (mothers being from 25 to 40 years old). Countries with larger positive coefficients in this eigenfunction are Bulgaria, Czech Republic, Slovakia, Hungary and U.S., while the Netherlands, Japan, Spain and Switzerland have negative coefficients.

The second eigenfunction (or eigensurface) reflects the specificity of the baby-boom around 1960 in Canada and U.S. (both have high positive coefficients in this eigenfunction). Countries with negative scores (such as Japan, Spain, Bulgaria, Hungary or Czech Republic) do not show a drop in fertility rates at the end of the 1960s. The third eigenfunction appears to correspond to a sudden drop at the end of the 1970s in fertility for women aged between 30 and 40 years. This could be associated with women's decision on reducing the number of children, as the high fertility rates for ages in the interval [30,40] before 1977 are mainly associated with large families or, in more technical terms, with high parities, *parity* being defined as the cumulative number of a woman's live births; see Preston *et al.* (2001)). This

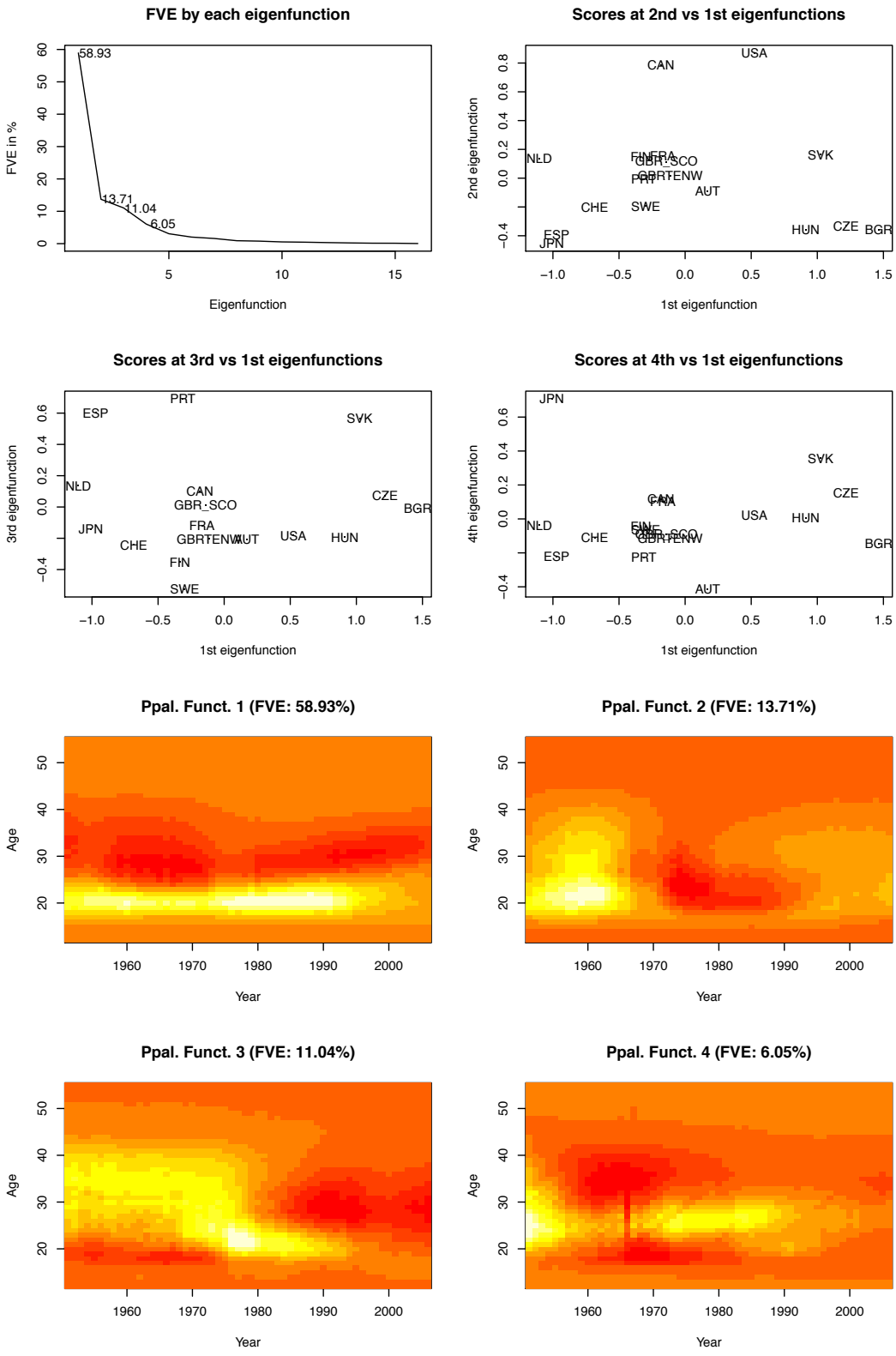

Figure 12: Standard FPCA of the fertility data $\text{ASFR}_i(s,t)$, $i = 1, \ldots, n = 17$, where the lower four panels display the first four eigensurfaces.

drop may be related to advances in birth control. These social changes arrived with a certain lag in countries with positive scores (Portugal, Spain, Slovakia) while they were adopted much earlier in countries with negative scores (Sweden, Finland, Switzerland). Other characteristics of this third eigenfunction are less intuitive.

Regarding the fourth eigenfunction, the score map in the panel in column 2, row 2, of Figure 12 indicates that Japan strongly weighs in this eigenfunction. Meanwhile the heat map (panel in column 2, row 4) shows a contrast between fertility concentrated around the age of 25 years (this strongly applies for Japan, with an outstanding positive score in this eigenfunction) and spread out fertility between the ages of 18 to 40, mainly between 1955 and 1980. Moreover, this heat map also shows an anomalous behavior (that appears as a discontinuity) at the year 1966. This fact corroborates that the fourth eigenfunction is a Japan specific function. We refer to the discussion in Section 5 for the anomaly in Japanese fertility in 1966.

Fitting the product FPC model to the fertility data resulted in estimates for the first four eigenfunctions ϕ_k of the operator $G_{\mathcal{T}}(t, v)$ as shown in Figure 13. The first of these time trend functions particularly weighs the pre-1990 fertility, while the others are contrasts between different calendar time periods. These estimated eigenfunctions are then multiplied with the age eigenfunction estimates $\hat{\psi}_j(s)$ of Figure 3 to obtain the product functions $\hat{\phi}_k(t)\hat{\psi}_j(s)$ that appear in the product FPC model representation (6) of ASFR(s, t). Figure 18 displays these product functions corresponding to the seven terms with larger FVEs among those with $j \leq 4$ and $k \leq 4$, which together explain 87.38% of the total variability; see also Table 1.

The product functions $\hat{\phi}_k(t)\hat{\psi}_j(s)$ in Figure 18 match well with the corresponding products $\hat{\phi}_{jk}(t)\hat{\psi}_j(s)$ in Figure 9 that result from the more general marginal approach (see Appendix B). These functions can thus be similarly interpreted as previously described in Section 5.1. The simplified product FPCA provides representations that are thus slightly less flexible and therefore explain somewhat less of the variance when compared with those

obtained from marginal FPCA, but have equally good, if not better, interpretability.

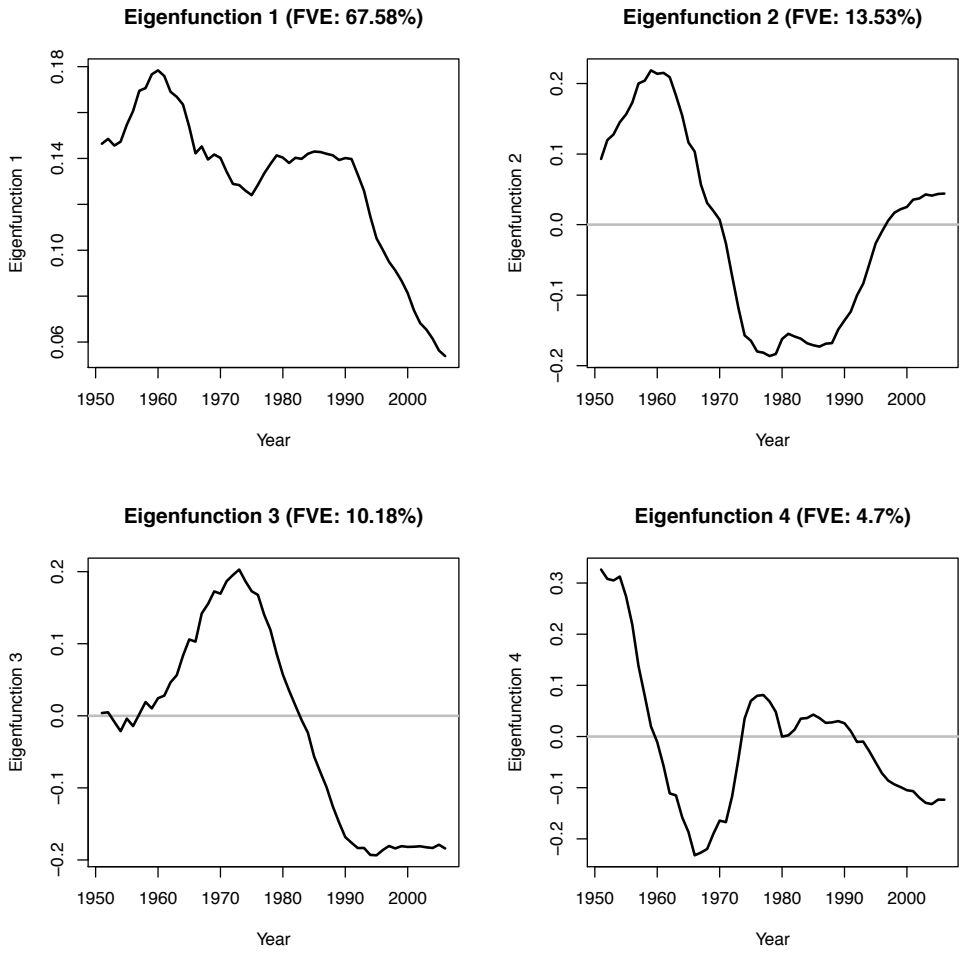

Figure 13: Estimated eigenfunctions $\hat{\phi}_k(t)$, $k = 1, 2, 3, 4$, in the product FPC model (6) for the fertility data.

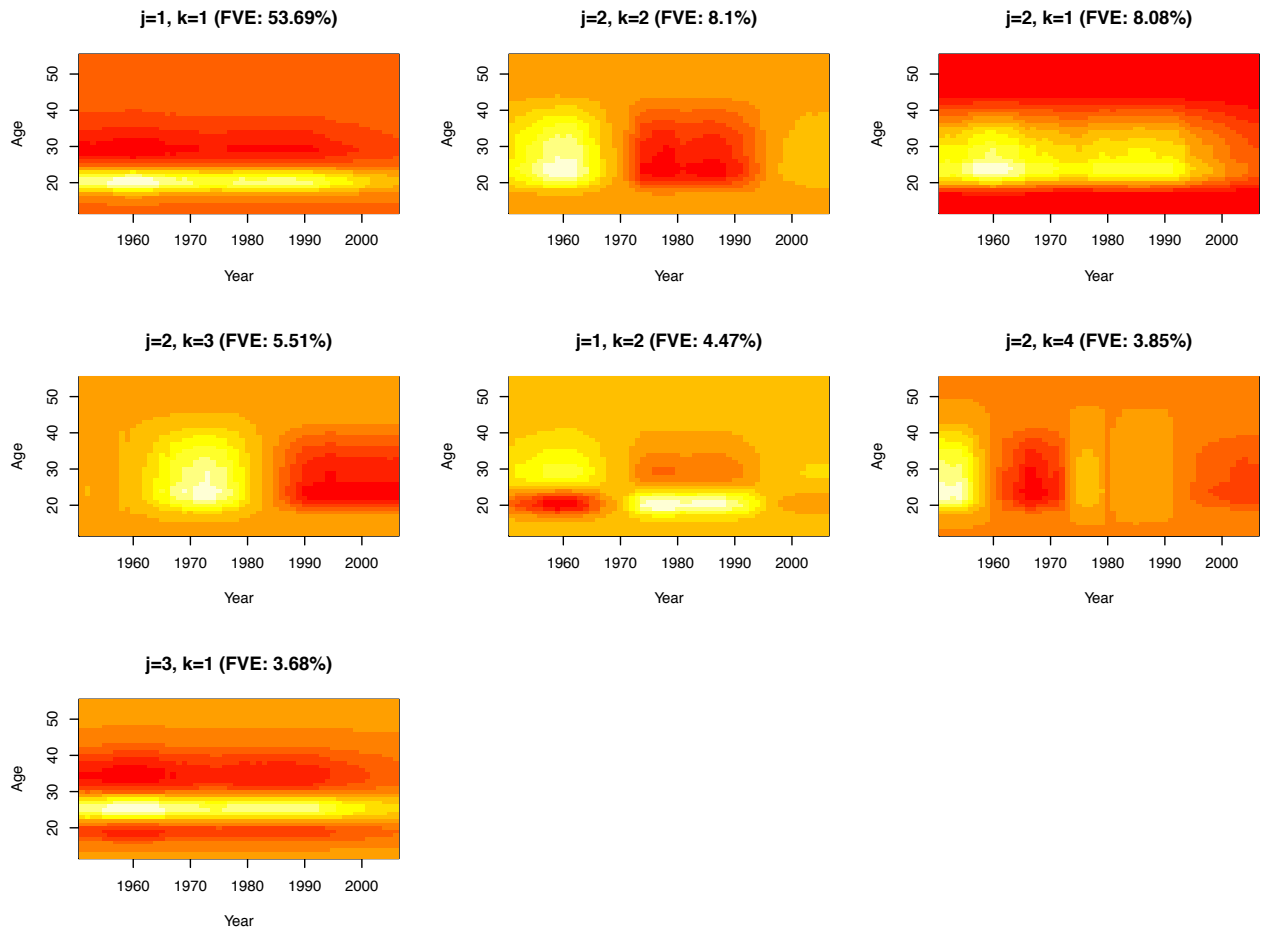

Figure 14: Product functions $\hat{\phi}_k(t)\hat{\psi}_j(s)$ corresponding to the seven terms with higher FVE in the product FPC model.

When applying product FPCA, one needs 7 terms to explain 87.38% of variance, while for the marginal FPCA it is sufficient to include 6 terms to explain 87.49% of the variance. Of course product FPCA has the big advantage that the final representation in general involves fewer functions $\hat{\psi}_j(s)$ and $\hat{\phi}_k(t)$ than the number of functions needed for the marginal FPCA representation and therefore is much simpler. For instance, the analysis of the fertility data with marginal FPCA involves 9 functions $(\hat{\psi}_1(s), \hat{\psi}_2(s), \hat{\psi}_3(s), \hat{\phi}_{11}(t), \hat{\phi}_{12}(t), \hat{\phi}_{21}(t), \hat{\phi}_{22}(t), \hat{\phi}_{23}(t), \hat{\phi}_{31}(t))$, while only 7 functions are involved in the product FPC model $(\hat{\psi}_1(s), \hat{\psi}_2(s), \hat{\psi}_3(s), \hat{\phi}_1(t), \hat{\phi}_2(t), \hat{\phi}_3(t), \hat{\phi}_4(t))$.

Supplement D: Male mortality rates as an additional example

Mortality rates (or death rates) are defined as a ratio of the death count for a given age-time interval divided by an estimate of the population exposed to the risk of death during some age-time interval (Preston *et al.* 2001). The Human Mortality Database (http://www.mortality.org/) provides detailed information on mortality rates for 37 countries or areas with precision of one year in both age and calendar time. Such rich information can be provided only by countries with well developed official statistics agencies. This is the reason why only 37 countries are covered by this database.

An alternative database including a much larger number of countries can be accessed through the Population Division of the Department of Economic and Social Affairs of the United Nations (WPP 2012). This database contains information for more than 200 countries on deaths grouped into five-year age intervals, from 1950 to 2010 (every 5 years). The price to be paid for including a much larger number of countries is a loss in precision, i.e., aggregation over 5 year intervals, both in terms of age and calendar time. As definition of the mortality rate for a given country during a period of consecutive years and an interval of ages, we consider the ratio between the number of deaths reported for a specific country over the selected 5 year calendar period for people with age at death in the selected 5 year age interval, divided by the number of people that at the beginning of the calendar time interval were in the age interval. As male and female mortality rates are different, we consider here male data that were downloaded (on the 14th of January 2015) from

```
http://esa.un.org/wpp/Excel-Data/EXCEL_FILES/3_Mortality/
                        WPP2012_MORT_F04_2_DEATHS_BY_AGE_MALE.XLS
http://esa.un.org/wpp/Excel-Data/EXCEL_FILES/1_Population/
                        WPP2012_POP_F15_2_ANNUAL_POPULATION_BY_AGE_MALE.XLS
```

We work with *log-Mortality Rates*, for which we use log(mortality rate+1), considered as functions of men's age grouped into intervals of 5 years (s) and repeatedly measured for

every 5 calendar years t for various countries. The aggregated log-mortalities constitute the functional data $X(s,t) = \text{log-mortality rate}(s,t)$.

In WPP (2012), data are provided for ages s in the year intervals $\{[0,5), [5,10), \ldots, [90,95), [95,\infty)\}$. The interval of calendar years with available data are $\{[1950,1955), \ldots, [2005,2010)\}$. The variability of mortality rates increases with age and decreases with population. So we limit ourselves to ages lower than 80. We also excluded countries with a 0 value for population size at any year or age. Then our database finally consisted of 166 countries, with 12 periods of five years each (which we labeled with the first year of the respective interval: 1950 to 2005) and 16 five years age intervals (labeled from 0 to 75).

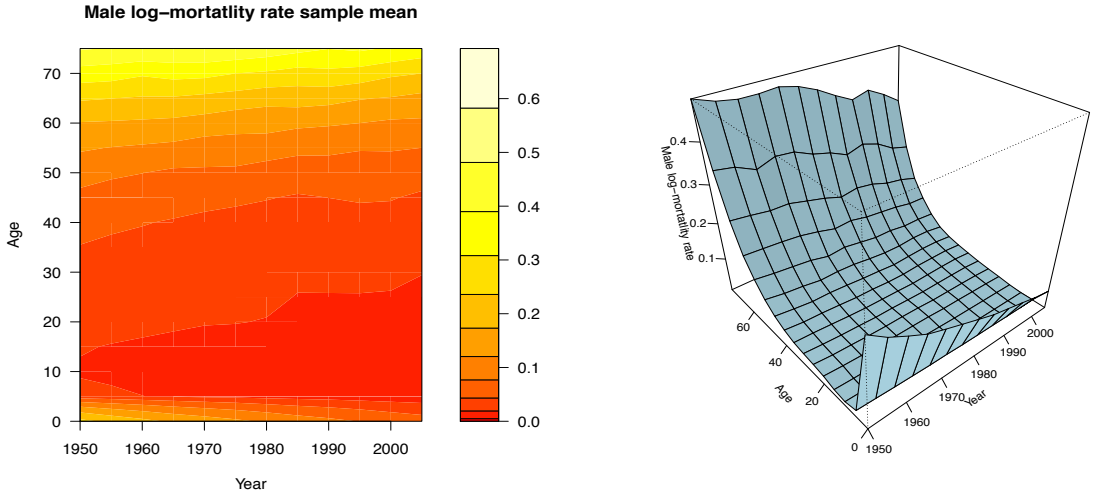

Figure 15: Sample mean of the 166 male log-mortality rate functions by calendar year.

The sample mean of the male log-mortality rate functions for 166 countries displayed in Figure 15 shows that mortality rates are, on average, highest for children under 5, and for men aged more than 60; and that they are decreasing with increasing calendar year.

The male log-mortality rate data include one log-mortality rate curve over age per calendar year and per country and are observed on a regular grid spaced in years across both coordinates age s and calendar year t, which means that the empirical estimators described in Section 2 can be applied to these data. Figure 16 displays the $nM = 1992$ centered functional data male log-mortality rates $X_i^c(s_l, t_m) = X_i(s_l, t_m)$ for $l = 1, \ldots, L = 16$,

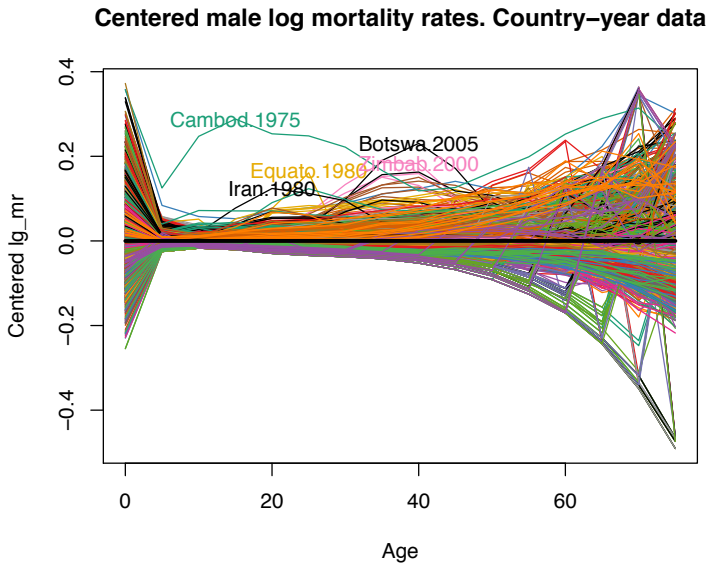

Figure 16: All available functional male log-mortality data as functions of *age* for 1992 combinations of 166 countries and 12 calendar years, centered around the mean. Functions corresponding to the same country are in the same color.

$m = 1, \ldots, M = 12$ and $i = 1, \ldots, n = 166$, demonstrating that there is substantial variation across countries and calendar years. Several outliers in the centered log mortality profiles have been highlighted in the figure. Some of these reflect periods of war, e.g. Iran 1980-1985 or genocides, e.g. Cambodia 1975-1980. Others correspond to high mortality rates due to the HIV/AIDS pandemic. The bloody reign of the Macias Nguema dictatorship in Equatorial Guinea also left its mark in this country's mortality profile.

We fitted the marginal FPC model and found that the $\phi_{jk}(t)$ are similar for $j = 1$ and 2. This is an indication that the product FPCA is appropriate for these data, and we directly applied it. Fitting the product FPC model to the male log-mortality data resulted in estimates for ψ_j and ϕ_k as shown in Figure 17. The shape of the first eigenfunction $\hat{\psi}_1(s)$ (that takes positive values for all ages s) is similar to that of the mean function for a fixed year t (see the right panel of Figure 15). Therefore $\hat{\psi}_1(s)$ can be interpreted as a *size* component: Country-years with positive score in the direction of this eigenfunction have higher male log-mortality ratios than the mean function for all ages, with larger differences

for larger values of the average log-mortality rates. The second eigenfunction $\hat{\psi}_2(s)$ represents a contrast between infant mortality and older age mortality. The third eigenfunction $\hat{\psi}_3(s)$ appears to point to difficulties in obtaining accurate estimates of mortality rates for the last age interval.

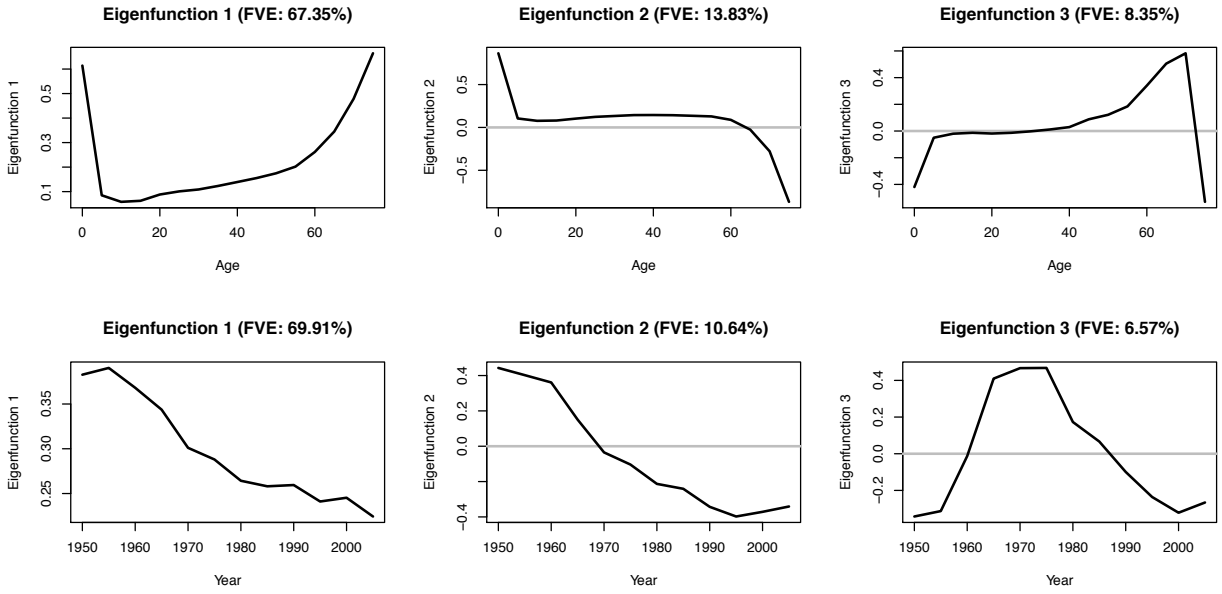

Figure 17: Estimated eigenfunctions $\hat{\psi}_j(s)$ (first row) and $\hat{\phi}_k(t)$ (second row), in the product FPC model for the log-mortality data.

The first calendar year trend function $\hat{\phi}_1(t)$ shows a continuous reduction in male log-mortality rates, with a pattern similar to the average evolution of male log-mortality over time (see Figure 15, right panel). So positive scores associated with this eigenfunction indicate larger reductions than the average (the opposite for negative scores). The second and third trend functions are contrasts between different calendar time periods. Positive (resp., negative) scores in the second trend function $\phi_2(s)$ indicate higher (resp., lower) than average mortality at the beginning of the overall calendar period, and lower than average mortality for the final years of the calendar period, i.e., a faster decline in mortality as compared to the average decline. The third eigenfunction is associated with differences in changes in log-mortality rates over calendar time between the middle period and the early/late periods.

The product functions $\hat{\phi}_k(t)\hat{\psi}_j(s)$ are shown in Figure 18. These functions can be easily interpreted by taking into account that they are the product of a function $\hat{\psi}_j(s)$ and a function $\hat{\phi}_k(t)$, as represented in Figure 17. When applying product FPCA, one needs 4 terms to achieve a FVE of 71.06%, and 6 terms to achieve a FVE of 75%.

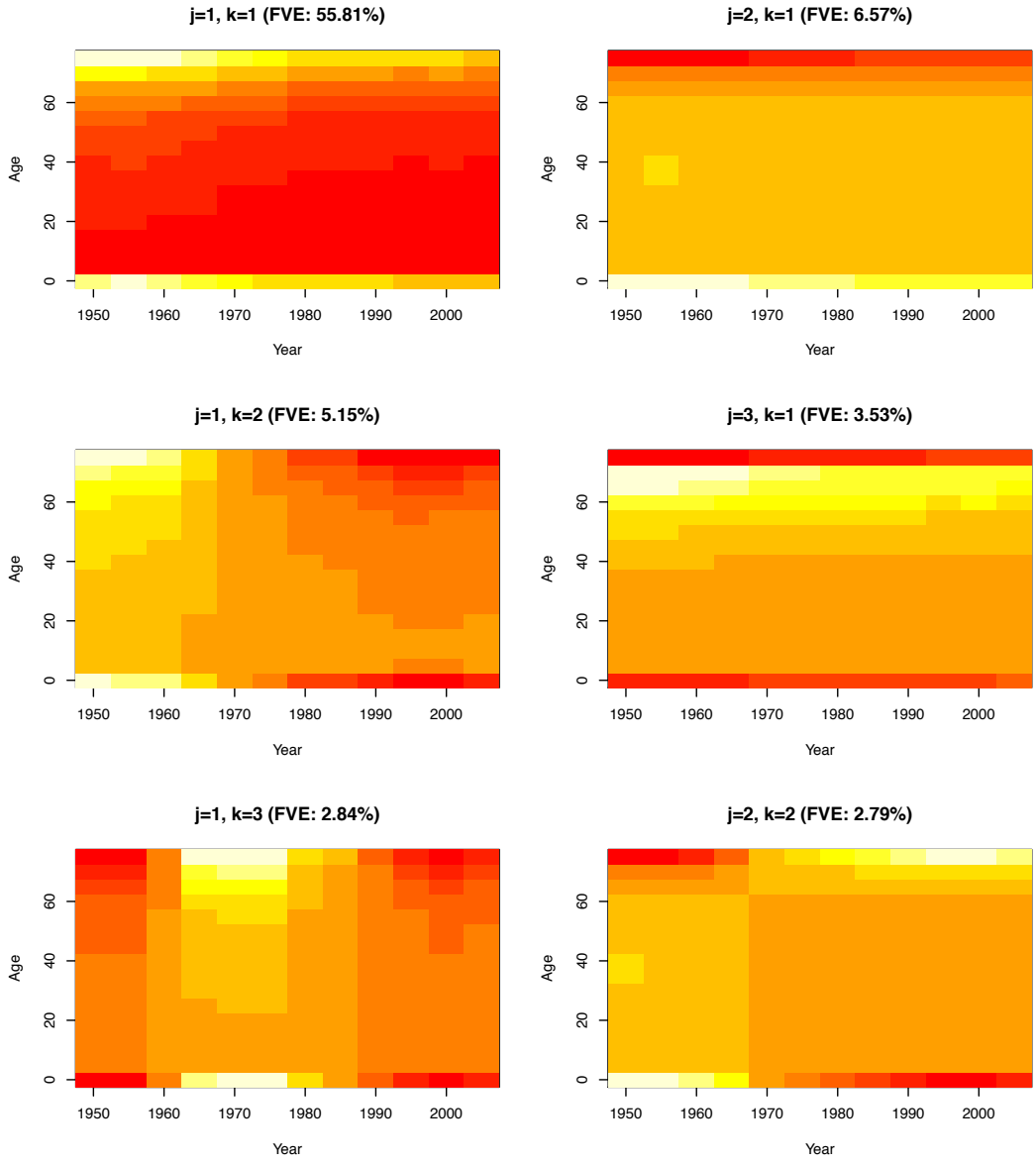

Figure 18: Product functions $\hat{\phi}_k(t)\hat{\psi}_j(s)$ corresponding to the six terms with higher FVE in the product FPC model representation for the log mortality data.

The first product of estimated eigenfunctions (with FVE equal to 55.81%) is $\hat{\phi}_1(t)\hat{\psi}_1(s)$, which is the product of the function $\hat{\phi}_1(t)$ that is similar to the average evolution of log-mortality rate over calendar years, and the function $\hat{\psi}_1(s)$ that has a shape similar to the average log-mortality rate pattern. As a consequence, the product function is always positive and very similar to the mean function (see Figure 15). So countries with positive random coefficients $\hat{\chi}_{11}$ at this product function $\hat{\phi}_1(t)\hat{\psi}_1(s)$ have larger male log-mortality rates than the average for all ages and all calendar years, with larger differences for larger values of the average log-mortality rates, and vice versa for the countries with a negative coefficient. We refer to Table 5 for a list of countries with most extreme (positive or negative) coefficients at this first product component.

The second product of eigenfunctions is $\hat{\phi}_1(t)\hat{\psi}_2(s)$ (FVE: 6.57%). It represents a contrast between infant mortality and old age mortality, due to the shape of $\hat{\psi}_2(s)$, which is more marked at the beginning than at the end of the period (because of $\hat{\phi}_1(t)$). Countries with negative scores (see Table 5) have lower than average infant log-mortality rates and higher than average old age log-mortality rates. The opposite applies to countries with positive scores at this product. The third product of eigenfunctions is $\hat{\phi}_2(t)\hat{\psi}_1(s)$ and it separates countries with faster than average reduction in male log-mortality rates (positive coefficients) from those with slower than average reduction (negative coefficients). This is the main effect of $\hat{\phi}_2(t)$. This effect is more marked for extreme ages, due to the shape of $\hat{\psi}_1(s)$. The countries with extreme coefficients as listed in Table 5 are extremes in a certain shape direction and deserve further study.

Alternatively, one can also apply marginal FPCA to quantify the observed variability across countries. The results for marginal FPCA are summarized in Figures 19 and 20 for the first three eigenfunctions, $\hat{\psi}_j(s)$, $j = 1, 2, 3$, resulting in a FVE of 89.52%. The first row of Figure 19 displays the estimated eigenfunctions $\psi_j(s)$, which are identical to the functions $\psi_j(s)$ used in the product FPCA. The second row of panels in Figure 19 shows the score functions $\hat{\xi}_{i,j}(t)$, $t \in \mathcal{T}$, which are country-specific functions of calendar year.

Table 5: Countries with the most extreme estimates of the random coefficients χ_{jk} obtained by fitting the product FPCA model (6) for the six terms with higher FVE in the representation of male log-mortality rates as linear combinations of the product functions $\hat{\phi}_k(t)\hat{\psi}_j(s)$.

	$\hat{\phi}_1(t)\hat{\psi}_1(s)$ (FVE: 55.81%)
Most −	Iceland, Channel Islands, Sweden, Norway, Puerto Rico, Barbados
Most +	Sierra Leone, Mali, Eritrea, Equatorial Guinea, Timor-Leste, Liberia
	$\hat{\phi}_1(t)\hat{\psi}_2(s)$ (FVE: 6.57%)
Most −	Fiji, Suriname, Martinique, Mauritius, Dem People's Republic of Korea, Guyana
Most +	Reunion, Central African Republic, El Salvador, Honduras, Pakistan, Angola
	$\hat{\phi}_2(t)\hat{\psi}_1(s)$ (FVE: 5.15%)
Most −	Channel Islands, Barbados, Iceland, Belarus, Rwanda, Sierra Leone
Most +	China, Oman, Tunisia, Singapore, Hong Kong SAR, Japan
	$\hat{\phi}_1(t)\hat{\psi}_3(s)$ (FVE: 3.53%)
Most −	Channel Islands, Iceland, Martinique, Guinea-Bissau, Timor-Leste, Oman
Most +	Reunion, Papua New Guinea, Eritrea, South Africa, Dem People's Republic of Korea, Guadeloupe
	$\hat{\phi}_3(t)\hat{\psi}_1(s)$ (FVE: 2.84%)
Most −	Cape Verde, Tajikistan, Kazakhstan, Azerbaijan, Belarus, Kyrgyzstan
Most +	Cambodia, Barbados, Channel Islands, Reunion, Guadeloupe, Martinique
	$\hat{\phi}_2(t)\hat{\psi}_2(s)$ (FVE: 2.79%)
Most −	Martinique, Japan, Fiji, Malta, Guyana, Botswana
Most +	Reunion, Barbados, Channel Islands, Iceland, Yemen, Eritrea

Their evolution over calendar time can be visualized by the track plot in Figure 20, showing the planar curves for the pairs $(\xi_{i,1}(t), \xi_{i,2}(t))$, $t \in \mathcal{T}$. In this example, the track plot is particularly useful to detect country-years with extreme scores in some eigenfunctions. For instance, Cambodia 1975-1980 and Rwanda 1990-1995 have extremely positive high scores in the first eigenfunction. This corresponds to periods in the history of these two countries during which they experienced a very high mortality rate: the Cambodian Genocide from 1975 to 1979, and the Rwandan Genocide in 1994.

The third step of the marginal FPCA (performing a separate standard FPCA for the estimated score functions $\hat{\xi}_{i,j}(t)$, $i = 1, \ldots, n$, for $j = 1, 2, 3$) yields estimated eigenfunctions $\hat{\phi}_{jk}$. For $k = 1, 2, 3$ these estimates are shown in Figure 19 (three lower rows). It can be seen that results are similar (up to sign changes) for the first and second sets of score functions.

To conclude this second example, we present the standard FPCA of the log-mortality data with the Karhunen-Loève representation, considering the data as random functions in two arguments. Figure 21 graphically summarizes the main results of this standard FPCA. The first four eigenfunctions have a FVE of 78.55%. There are similarities between these eigenfunctions and, respectively, the 1st, 2nd, 3rd and 5th eigenfunction products represented in Figure 18 (in the two last cases, up to a sign change). Therefore the interpretation we have made above for these eigenfunctions products are valid for the eigenfunctions obtained by standard FPCA. Nevertheless, to arrive at these interpretations is much more difficult if the starting point is Figure 21, without the benefit of the functions represented in Figure 18 for the product FPCA and their decomposition as products of functions in Figure 17.

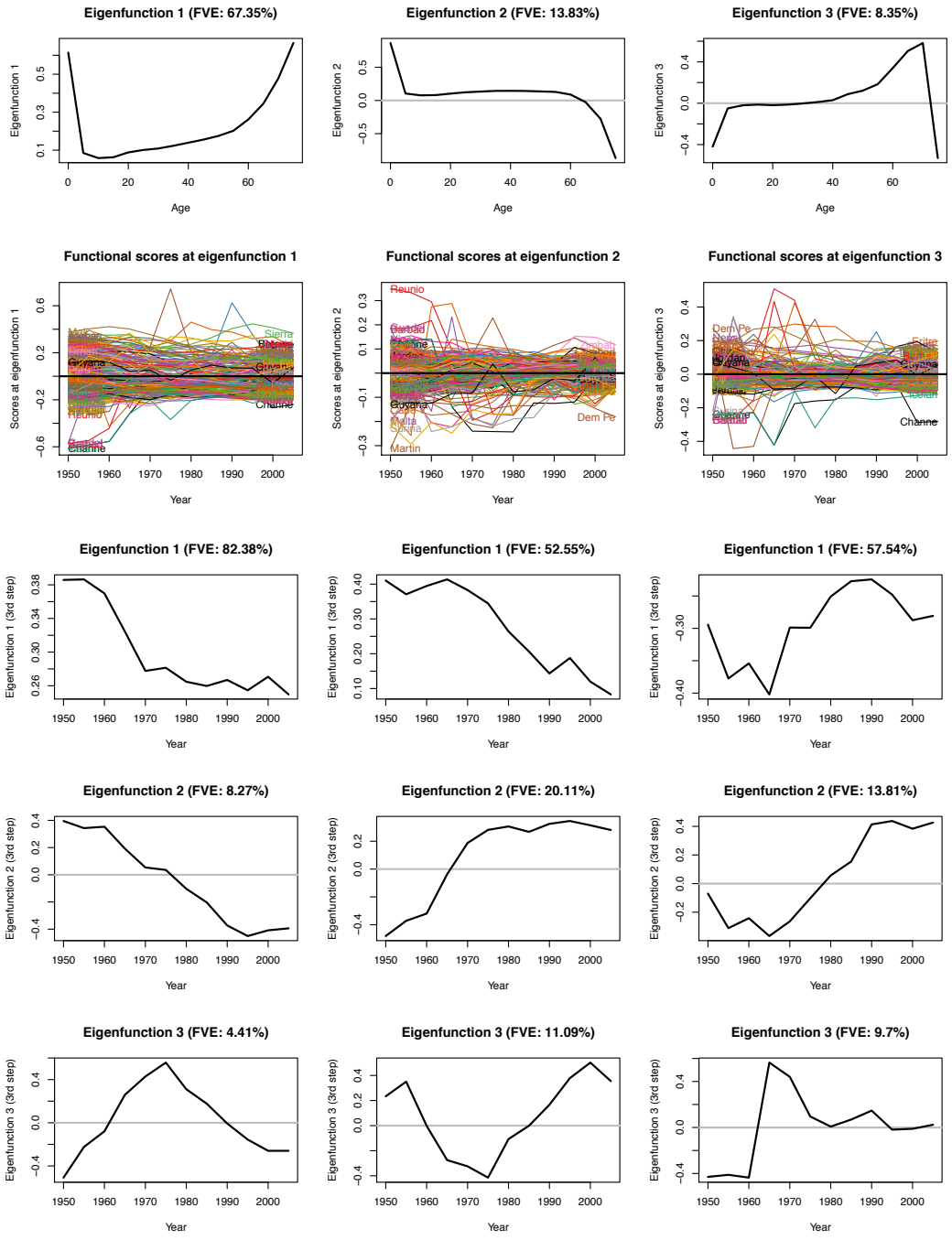

Figure 19: Results of the marginal FPCA for the male log-mortality rate data. Columns 1, 2 and 3 correspond to eigenfunctions 1, 2 and 3 in the second step of the marginal FPCA, respectively. *First row:* Estimated eigenfunctions $\hat{\psi}_j(s)$, where s is age. *Second row:* Score functions $\hat{\xi}_{i,j}(t)$, where t is calendar year. *Rows 3, 4 and 5:* Estimated eigenfunctions $\hat{\phi}_{jk}(t)$, $k = 1, 2, 3$, in the third step.

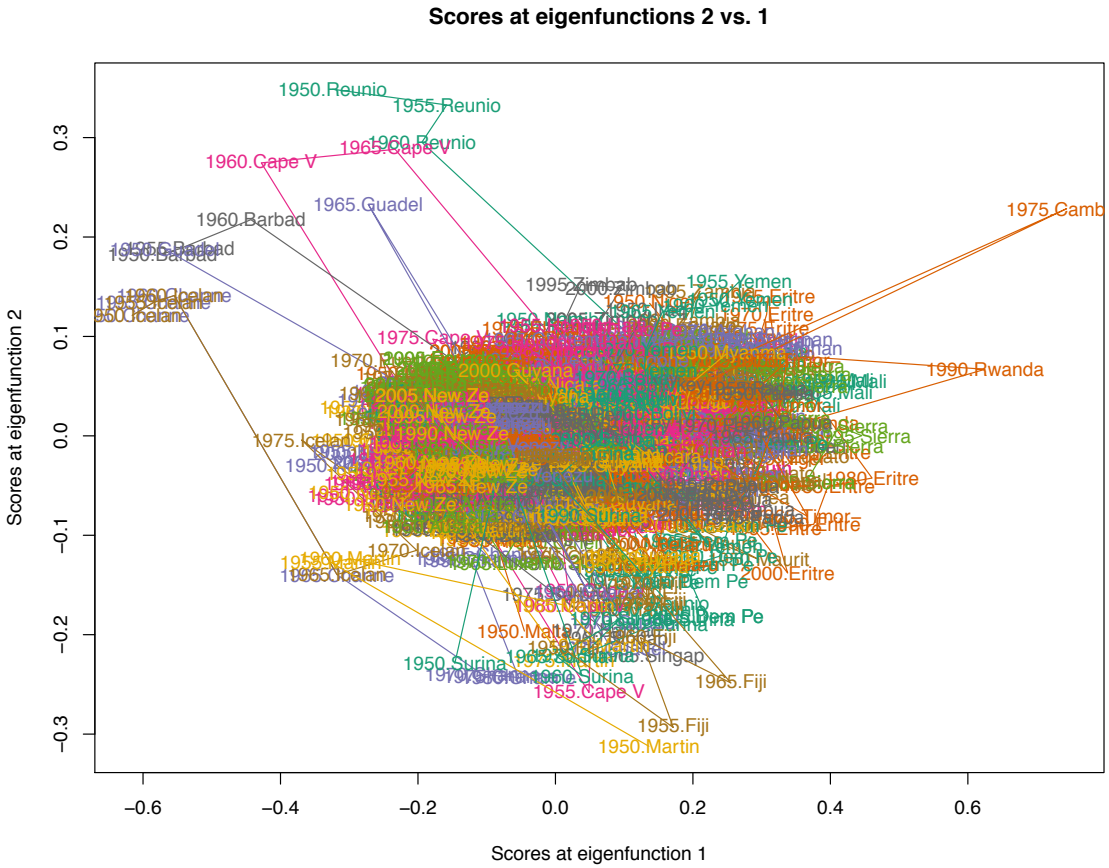

Figure 20: Track-plot corresponding to the implicitly parametrized planar curves $\{(\hat{\xi}_{i,1}(t), \hat{\xi}_{i,2}(t)) : t = 1950, 1955, \ldots, 2005\}$, parametrized by calendar time t, where $\xi_{i,j}(t)$ is the j-th score function for country i.

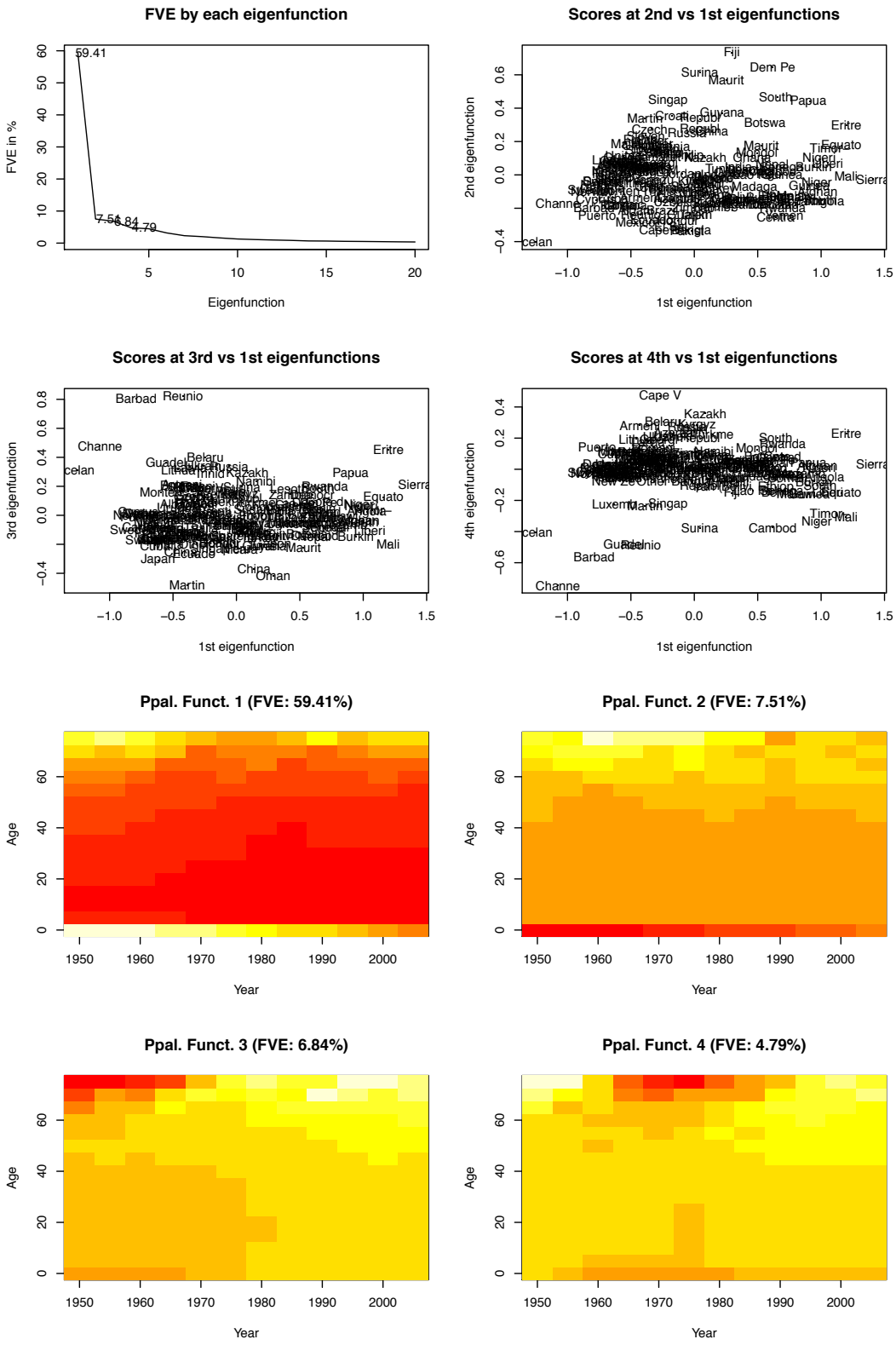

Figure 21: Standard FPCA of the male log-mortality data.

REFERENCES

Bosq, D. (2000) *Linear Processes in Function Spaces: Theory and Applications*. New York: Springer.

Horváth, L. and Kokoszka, P. (2012) *Inference for Functional Data with Applications*. New York: Springer.

Preston, S. H., Heuveline, P. and Guillot, M. (2001) *Demography: Measuring and modeling population processes*. Blackwell Publishing.

WPP (2012) World population prospects: The 2012 revision, dvd edition.

Yao, F., Müller, H.-G. and Wang, J.-L. (2005) Functional data analysis for sparse longitudinal data. *Journal of the American Statistical Association*, **100**, 577–590.